

Using Transparency Models to evaluate Open Data Systems: a case study using a Mexican Open Data Platform

Hugo Estrada, Blanca Vazquez, and Karen Najera

Fondo de Información y Documentación para la Industria - INFOTEC, Mexico
{hugo.estrada, blanca.vazquez, karen.najera}@infotec.com.mx

Abstract. Actualmente, la transparencia es un tema demandado por la sociedad principalmente en el ámbito político. En este sentido, muchos organismos gubernamentales han buscado transparentar su información mediante el uso de plataformas de software que exponen públicamente sus datos. El uso de éstas plataformas se ha generalizado en muchos países, y es posible evaluar el nivel de apertura de los datos que publican. Sin embargo, no es posible determinar su nivel de transparencia como sistemas de software. En este artículo, proponemos el uso de modelos de transparencia desarrollados por expertos para evaluar cualitativamente el nivel de transparencia de una plataforma de Datos Abiertos desarrollada en México. Como caso de estudio presentamos los resultados previos de ésta evaluación los cuales comprenden el nivel de transparencia de la plataforma como sistema de software así como el nivel de transparencia de la información que publica.

Keywords: Transparencia, Datos Abiertos, Modelo de Transparencia, Plataforma de Datos Abiertos

1 Introducción

Hoy en día, la transparencia es un tema demandado por la sociedad, la cual solicita a los gobiernos e instituciones privadas hacer pública la información relativa a sus procesos. En este sentido, la iniciativa de Datos del Gobierno Abierto (OGD, por sus siglas en inglés) promueve la transparencia mediante la apertura de los datos gubernamentales, impulsando a la ciudadanía hacia la consulta, uso y valoración de la información [2]. La OGD abarca los conceptos de Datos Abiertos (aquellos que son distribuidos libremente por cualquier persona) y Datos abiertos del gobierno (aquellos que son producidos por organismos públicos). Para que los Datos de Gobierno puedan ser considerados Datos Abiertos deben cumplir con los siguientes principios (propuestos por la Fundación Sunlight¹ y el grupo de trabajo en Gobierno Abierto²): ser completos, sin previo procesamiento, oportunos, accesibles, procesables, no discriminatorios, descritos en formatos no propietarios y publicados bajo licencia abierta. El nivel de apertura de los Datos de Gobierno puede ser medido mediante el esquema de apertura

¹ Sunlight Foundation. <http://sunlightfoundation.com/>

² World Wide Web Consortium (W3C). <http://www.w3.org/Consortium/>

de Datos Abiertos propuesto por la W3C [3], que define que la apertura de los datos es mayor mientras más alto sea el nivel en el que se encuentran. Los niveles son: *1 estrella*, datos disponibles en la Web con licencia abierta, en cualquier formato (e.g. PDF); *2 estrellas*, datos en formato estructurado (e.g. Ms Excel); *3 estrellas*, datos en formato estructurado y no propietario (e.g. CSV, XML); *4 estrellas*, datos en formato estructurado, no propietario, que usan Identificadores Uniformes de Recursos (URIs) y estándares de la W3C (e.g. RDF) para identificar cosas y propiedades; *5 estrellas*, datos cuyas URIs apunten a otros datos dotándolos de contexto (datos ligados).

Actualmente, se han desarrollado diversas plataformas para la publicación de Datos Abiertos, cuyo uso se ha generalizado en los gobiernos de muchos países. Si bien es posible analizar el nivel de apertura de los datos utilizando el esquema de la W3C, no es posible determinar el nivel de transparencia de estas plataformas como sistemas de software y debido a su propósito, estas plataformas deben proveer un alto nivel de transparencia. Este tema ha sido abordado por Leite [4], quien afirma que la transparencia del software se está convirtiendo en un criterio de calidad relevante, por ello, hace uso de modelos definidos por Capelli [1] (Modelo de Transparencia y Modelos de Operacionalización), que definen la transparencia de procesos (desde el punto de vista cualitativo), para proporcionar una definición de transparencia desde la perspectiva de software.

En este artículo se propone el uso del Modelo de Transparencia y de los Modelos de Operacionalización [1, 4], como medio para evaluar el nivel de transparencia de sistemas de software. Como caso de estudio se utiliza una plataforma de Datos Abiertos desarrollada por el Fondo de Información y Documentación para la Industria INFOTEC en México, de la cual se estima su nivel de transparencia como sistema de software así como su nivel de transparencia como sistema de información y se presentan los resultados previos de la evaluación de la plataforma.

2 Objetivos de investigación

El objetivo de este artículo es demostrar cómo el Modelo de Transparencia y los Modelos de Operacionalización [1, 4] pueden ser utilizados para evaluar de manera cualitativa el nivel de transparencia de plataformas de Datos Abiertos. En este artículo se evalúa una plataforma de Datos Abiertos desarrollada en INFOTEC, México, analizándola como sistema de software y como sistema de información.

3 Contribuciones esperadas

Las contribuciones esperadas de este trabajo a partir de los resultados obtenidos de la evaluación son:

- Un caso de éxito que demuestre que el Modelo de Transparencia y los Modelos de Operacionalización [1, 4] pueden utilizarse para evaluar el nivel de transparencia de sistemas de software orientados al acceso público de información, como las plataformas de Datos Abiertos y los portales de transparencia.

4 Resultados alcanzados

La plataforma de Datos Abiertos desarrollada por INFOTEC integra un conjunto de funcionalidades para la publicación, gestión y consumo de datos públicos, con el objetivo de proveer a los ciudadanos un espacio para la participación y colaboración. Al ser una plataforma que promueve la transparencia de información, consideramos que es necesario conocer su nivel de transparencia e identificar sus puntos de mejora. Por este motivo proponemos utilizar el Modelo de Transparencia y los Modelos de Operacionalización propuestos en [1, 4] para evaluar nuestra plataforma.

Nuestra evaluación se basó en el trabajo de Chung [5], el cual consiste en el análisis de las funcionalidades de un sistema y su mapeo contra aquellos requisitos no funcionales que las describen. El uso de la investigación de Chung fue crucial para evaluar cualitativamente la plataforma, ya que el Modelo de Transparencia y los Modelos de Operacionalización describen atributos de calidad (conocidos como requisitos no funcionales).

Previo a la evaluación de la plataforma se estudió el Modelo de Transparencia, identificando las definiciones de cada una de sus características (*accesibilidad, usabilidad, informatividad, entedimiento y auditabilidad*) y sus correspondientes sub-características (la tabla 1 presenta *usabilidad* e *informatividad* y sus sub-características). Cada sub-característica cuenta con un Modelo de Operacionalización que incluye elementos que detallan las acciones que la sub-característica debe cumplir. Estos Modelos de Operacionalización también fueron estudiados. El proceso de evaluación es el siguiente:

Para determinar si la plataforma cubría con una característica del Modelo de Transparencia, se realizó una evaluación de abajo hacia arriba, partiendo del Modelo de Operacionalización hacia las características del Modelo de Transparencia. La evaluación consistió en el mapeo manual de las funcionalidades de la plataforma, con cada uno de los elementos de los Modelos de Operacionalización. Si al menos uno de los elementos de un Modelo de Operacionalización era cubierto por las funcionalidades de la plataforma, entonces se determinaba que la sub-característica era cubierta. No se exigió que todos los elementos fueran cubiertos por la plataforma dado que algunos de ellos están orientados a procesos y las funcionalidades de la plataforma no están enfocadas a procesos.

Una vez determinadas las sub-características cubiertas por la plataforma, se definió que si todas las sub-características pertenecientes a una característica eran cubiertas, entonces la característica era cubierta totalmente, y si no todas pero al menos una de las sub-características pertenecientes a una característica era cubierta, entonces la característica era cubierta parcialmente.

A continuación se describen los resultados previos de la evaluación. La Fig. 1 muestra las características cubiertas por nuestra plataforma de Datos Abiertos.

4.1 Transparencia desde la perspectiva de sistema de software

Desde la perspectiva de sistema de software, la plataforma de Datos Abiertos satisface la característica base de *Usabilidad* del modelo de transparencia, debido a que cubre, en forma satisfactoria todas sus sub-características. A continuación se describe cómo son cubiertas las sub-características por la plataforma.

Table 1. Definición de características del modelo de transparencia [1, 4]

Característica	Definición
Usabilidad	Cualidad de ser capaz de proporcionar un buen servicio.
Uniformidad	Cualidad de ausencia de diversidad y variación.
Simplicidad	Cualidad de estar libre de dificultades y esfuerzos.
Operabilidad	Cualidad de ser operacional.
Intuitividad	Cualidad de ser espontáneo derivado o provocado por una tendencia natural.
Desempeño	Capacidad de dar buen rendimiento.
Adaptabilidad	Capacidad de cambiar o ser cambiado de acuerdo a las circunstancias.
Amigable	Capacidad para usar fácilmente.
Informatividad	Capacidad de proveer o cubrir información.
Claridad	Capacidad de ser libre y fácil de entender.
Completo	Cualidad de estar completo y entero, teniendo cada cosa que se necesita.
Exactitud	Capacidad de ser libre de errores.
Actual	Cualidad de ocurrencia o pertenencia en tiempo presente.
Comparable	Capacidad para ser comparado.
Consistencia	Capacidad de obtener resultados similares en cada ejecución.
Integridad	Cualidad de integridad indivisible, intactos o totalidad.
Precisión	Capacidad de precisión sin errores sistemáticos.

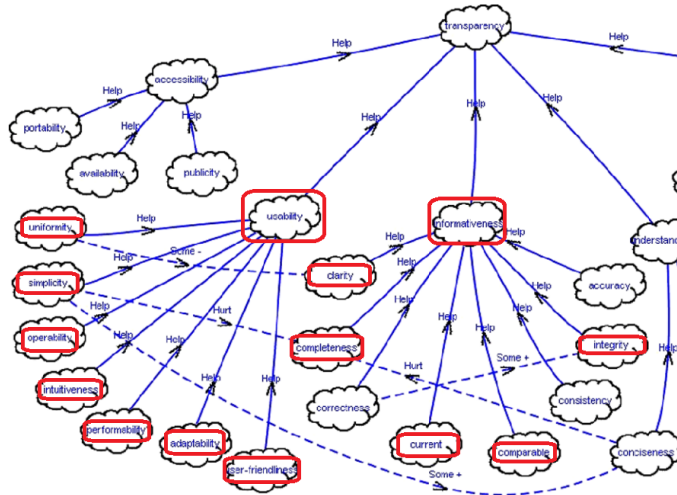


Fig. 1. Características del Modelo de Transparencia [4] que cubre la plataforma

Uniformidad: La plataforma es uniforme debido a que cuenta con un diseño que aplica un mismo patrón visual a toda la plataforma. También cuenta con un patrón e imagen visual definida para categorizar cada uno de los tipos de archivos que se almacenan. De esta forma, se da el mismo tratamiento a todos los tipos de archivos, ayudando a uniformizar su acceso y utilización.

Simplicidad: La plataforma es simple debido a que cuenta con rutas bien definidas de navegación para que el usuario acceda a datasets y aplicaciones, así como se presenta una unificación de metadatos para el registro de estos y de usuarios.

Además únicamente archivos con formatos no propietarios pueden ser almacenados en la plataforma (CSV, KML, KMZ, XML y JSON).

Operabilidad: La plataforma es operable debido a que cualquier usuario puede acceder a todas las funcionalidades, descargar o visualizar información sin restricciones y sin requerir previa capacitación.

Intuitividad: La plataforma es intuitiva debido a que se utilizan patrones bien definidos para interactuar con las funcionalidades relativas a los datasets y aplicaciones. Estos patrones incluyen imágenes y texto que ayudan al usuario a intuir fácilmente en qué consiste cada funcionalidad y si le es de utilidad o no.

Desempeño: La plataforma cubre esta característica debido a que cuenta con módulos específicos para mejorar el desempeño en cuanto al almacenamiento y acceso a los datos en situaciones concurrentes, esto se logra a través del mejoramiento de conectores para acceso a bases de datos. Además esto es soportado por servidores dedicados, lo cual permite que la plataforma pueda almacenar grandes cantidades de información con un alto desempeño.

Adaptabilidad: La plataforma es adaptable debido a que se construyó sobre una plataforma de desarrollo dirigida por ontologías [6], por lo que una ontología es la base de la plataforma. La ontología define la estructura del sistema, el modelo de datos y el comportamiento de objetos, así también esta puede ser fácilmente extendida para soportar nuevos requerimientos.

Amigable: La plataforma es amigable, ya que cumple con las características de *uniformidad*, *simplicidad*, *operabilidad* e *intuitividad*, lo cual hace que la plataforma pueda ser usada fácilmente.

4.2 Transparencia desde la perspectiva de información

Desde la perspectiva de la información que proporciona, la plataforma de Datos Abiertos satisface parcialmente la característica base de capacidad *Informativa* del modelo de transparencia, debido a que cubre, en forma satisfactoria la mayoría de sus sub-características. A continuación se describe cómo son cubiertas las sub-características por la plataforma.

Completa: La información proporcionada en la plataforma es completa debido a que se definió un conjunto de metadatos que describen a detalle datasets, aplicaciones, usuarios consumidores y publicadores.

Claridad: La información proporcionada en la plataforma es clara debido a los metadatos mencionados en el punto anterior. Además se definió una guía de publicación que describe estos metadatos, y los formatos de archivo permitidos.

Integridad: La información proporcionada en la plataforma es íntegra, ya que los datos proporcionados por los publicadores se almacenan sin modificaciones ni alteraciones, garantizando así la integridad de los datos.

Comparable: La información proporcionada en la plataforma es comparable debido a que se proporcionan estadísticas del uso de datos, aplicaciones, de comentarios, de descargas, de visualizaciones y de puntuaciones. Las estadísticas pueden ser visualizadas en forma tabular o gráfica.

La sub-característica *exactitud* no se cubre ya que la información expuesta en la plataforma es introducida por los usuarios, lo que no garantiza que sea libre de errores. Las sub-características *consistencia* y *precisión* engloban acciones de ejecución, con las que se podría evaluar como sistema de software a la plataforma, sin embargo no aplican para evaluar la información publicada en ésta.

5 Conclusiones

En el presente artículo se describió un caso de estudio en el que se evaluó el nivel de transparencia de una plataforma de Datos Abiertos desarrollada en México mediante el uso de modelos de transparencia definidos por expertos en el dominio. El caso de estudio fue realizado con la finalidad de demostrar que los modelos de transparencia pueden utilizarse para evaluar el nivel de transparencia de sistemas de software orientados al acceso público de información.

Como resultado del caso de estudio se identificó la necesidad de proporcionar una definición genérica de las características del modelo de transparencia y sus correspondientes modelos de operacionalización, debido a que las definiciones actuales están enfocadas a procesos. De tal manera, un modelo genérico sería de gran utilidad para evaluar la transparencia de un sistema de software desde diferentes perspectivas, por ejemplo, la transparencia como sistema de software y la transparencia de la información que expone.

Como trabajo futuro, con base en los resultados del caso de estudio, se plantea la definición de las funcionalidades básicas que un sistema de software orientado al acceso público de información debe considerar para transparentar el software y la apertura de su información.

Es importante comentar que la plataforma analizada en este artículo es un producto comercializable que comprende también sus servicios asociados como consultoría, soporte técnico, capacitación y personalización.

References

1. Claudia Cappelli Aló. *Uma Abordagem para Transparência em Processos Organizacionais Utilizando Aspectos*. PhD thesis, Pontifícia Universidade Católica Do Rio de Janeiro, 2009.
2. Ubaldi. B. Open government data: Towards empirical analysis of open government data initiatives. *OECD Working Papers on Public Governance*, (22):127–139, 2013.
3. Tim Berners-Lee. Linked Data. <http://www.w3.org/DesignIssues/LinkedData.html>, 2009. Last access: 23/08/2013.
4. Julio Cesar Sampaio do Prado Leite and Claudia Cappelli. Software transparency. *Business & Information Systems Engineering*, 2(3):127–139, 2010.
5. John Mylopoulos, Lawrence Chung, and Brian A. Nixon. Representing and using nonfunctional requirements: A process-oriented approach. *IEEE Trans. Software Eng.*, 18(6):483–497, 1992.
6. Javier Solis, Hasdai Pacheco, Karen Najera, and Hugo Estrada. A MDE Framework for Semi-automatic Development of Web Applications. In *Proceedings of the 1st International Conference on MDE and Software Development*, pages 241–246, 2013.